DOCUMENT RESUME

ED 083 265 TM 003 226

AUTHOR Fincher, Cameron
TITLE Program Evaluation: Approaches and Procedures.
PUB DATE 16 May 73
NOTE 18p.; Paper presented at annual forum of Association for Institutional Research (Vancouver, British Columbia, May 16, 1973)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Conceptual Schemes; *Evaluation Techniques; *Higher Education; *Program Evaluation; Research Reviews (Publications); Speeches

ABSTRACT

Current approaches being made to program evaluation are examined, and some of the implications for academic programs in higher education are considered. The conceptual shifts that are involved in the changing demands placed upon educational evaluation are summarized as: (1) an increasing emphasis on the use of tests and measurements for feedback and direct improvement as opposed to selection and placement of students, (2) an increasing advocacy of developmental, applied, practical, problem-solving research, (3) a lessening of interest in conventional experimental research as opposed to research that is administratively or action oriented, and (4) a growing interest in decision-oriented research as opposed to conclusion-oriented research. For the evaluation of academic programs, these shifts suggest that a heavier emphasis should be given to formative, as opposed to summative, evaluation, and that the results of evaluation studies will be used increasingly for policy decisions that will determine program substance, form, and content. A conceptual framework for program evaluation in higher education is provided in a figure. The primary thrust of the conceptual design is that systematic efforts at program evaluation should consitute a functional analysis of the variables that are involved, and that the nature of the relationships derived or established will be probabilistic rather than deterministic. The complexities of most higher education program and the subtlety of influenctial variables suggest that urgent as evaluation is, it will not be easy to conduct the kind of evaluation studies needed. (DB)

FILMED FROM BEST AVAILABLE COPY

ED 083265

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PROGRAM EVALUATION: APPROACHES AND PROCEDURES

Cameron Fincher
Institute of Higher Education
University of Georgia

TM 003 226

Paper presented at the annual forum of the Association for Institutional Research in Vancouver, British Columbia, May 16, 1973

PROGRAM EVALUATION: APPROACHES AND PROCEDURES

Cameron Fincher
Institute of Higher Education
University of Georgia

There is an increasing concern with program evaluation in higher education. This concern stems from a demand for new academic programs to meet the changing needs of college students and from an insistence that programs be properly managed and systematically judged.

The Carnegie Commission on Higher Education (1972) has issued a strong appeal for the development of academic programs that would serve the needs of students who have not traditionally sought higher education and whose motives and interests differ greatly from previous students. The clear implication of that appeal is that students are changing and academic programs must be changed accordingly. The Commission on Nontraditional Study (1973) makes an even stronger appeal for academic programs that would serve students rather than institutions, permit a degree of flexibility and an adaptiveness that has often been advocated but seldom realized, and help realize a total learning society that would place no limits on the efforts of its citizens to acquire the skills and knowledge they seek.

Other converging events that spur the concern with program evaluation are the federal legislation of the sixties that made specific requirements for evaluation and the current dominance of economic and engineering concepts in educational thought and discussion. The Elementary and Secondary Education Act of 1965 (ESEA) gave an impetus to evaluation that may indeed have a greater impact on education than the act itself. The allocation of massive sums for the improvement of education did carry a mandate for insuring that the money was well spent. The ensuing disillusionment with the massive

programs funded by ESEA underscored the need for better methods, skills, and techniques for the evaluation of educational programs. To quote Malcolm Provus:

> Perhaps before we can build effective new programs, we must establish creative new ways to monitor and eventually judge the effectiveness of such programs. This capacity to evaluate programs must ultimately depend upon a management theory that utilizes pertinent, reliable information as the basis for administrative decisions (1969, p. 243).

The dominance of economic and engineering concepts has emphasized the shift of conceptual focus from input and process variables to an assessment of educational outcomes. The overplayed theme of accountability is based on a concern with efficiency in the utilization of personnel, resources, and materials in the educational enterprise. As much as educators may disclaim the appropriateness of economic and engineering concepts for academic programs, they should be nonetheless cognizant that such concepts are the dominant ones at the present time.

The purpose of this paper is to examine the current approaches that are being made to program evaluation and to consider some of the implications for academic programs in higher education. The effort is predicated on the basis that program evaluation is a management imperative for higher education and vitally affects the role and function of institutional research in our colleges and universities. The evaluation of academic programs in the past has not been viewed within the larger framework of policy analysis and has not been a major concern of institutional research as such. Not only have past efforts been spasmodic and piecemeal, but the evaluation of academic programs has not been well grounded and has often produced results that conflict with professional experience and judgment. The Coleman Report (1966), the study by the RAND Corporation for the President's Commission on School

Finance (1972), and the much publicized Jencks study (1972) are indicative of an increasingly persuasive argument that runs counter to our better hopes and expectations for education. In brief, these studies raise a number of critical questions for public education as it is traditionally valued in American society. They leave us wondering most seriously where the difficulty lies -- in the educational programs as implied, in the methodology of evaluation as it is imperfectly developed, or in the choice of criteria as they are opportunistically employed. The serious implications of such studies strongly suggest that they cannot be easily ignored -- and they cannot be accepted at face value.

The dilemma posed by the numerous efforts at program evaluation in the past several years is, to say the least, embarrassing. It would seem most advisable then to examine recent trends in program evaluation, to understand carefully the conceptual and methodological difficulties that are involved, and to seek a better conceptual grasp of the nature and functions of evaluation in general.

RECENT TRENDS IN EDUCATIONAL EVALUATION

There is a well-known but highly embarrassing suspicion that most efforts to evaluate human productivity, effectiveness, or well-being wind up on a pessimistic note. This is true not only of programs in education but also of evaluative efforts in the fields of social welfare, penal corrections, public health, counseling, psychotherapy, rehabilitation, public housing, and religious conversion. Where traditional methods of experimental or quasi-experimental investigation have been applied to the problems of human betterment, correction, enlightenment, or uplift, the results have not been as encouraging as proponents of the program, technique, procedure, cure, or

remedy had hoped in the beginning. Whatever human nature is, it does not appear as malleable, flexible, or plastic as reformers, revivalists, and crusaders proclaim. Change in human behavior is not easily assessed in systematic, quantitative, or experimental fashion and an inherent resistance to change has been postulated more than once to avoid the full pessimism of research findings.

Perhaps the encouraging feature of the most recent wave of pessimism is the growing recognition that the scientific methods developed for investigation of physical and biological phenomena in a laboratory setting are not adequate for the evaluation of programs seeking to improve, correct, or benefit human beings. It is not the wrongness of scientific method that is underscored by this outlook but rather the limitations of narrowly construed methods in dealing with social, cultural, and psychological complexities.

A CHANGING NOTION OF EVALUATION

As in so many other changing trends in education, the National Society for the Study of Education was quick to address one of its yearbooks to a consideration of the issue and problems involved. Edited by Ralph Tyler (1969) the 1969 yearbook addresses itself to the changes taking place in educational evaluation. New needs, conditions, knowledge, and technologies are said to foster new concepts, procedures, and instruments of evaluation. Massive financial support was first given by the federal government to the development of new courses in science and mathematics, then to the social problems of poverty and discrimination. Each produced in turn a demand for sound data that would guide public decision and policy. Coupled with the rapid increase in technological instrumentation, federal support of education spurred a concern for evaluation that could not be met by traditional test

theory and practices. An adequate appraisal of new conditions and the application of new knowledge about education as such was believed to necessitate new developments in evaluation.

Merwin (1969) emphasized that change had been prevalent in educational evaluation since its inception in the 19th century but changing concepts of evaluation were called for by the new educational efforts and technical developments. Bloom (1969) found one of the major theoretical issues of evaluation to be the degree of specification of educational objectives and outcomes. The basic problem is making the desired outcomes of learning sufficiently explicit for communication to teachers, technicians, students, and the evaluators themselves. There is also the problem of evaluating the non-specified outcomes of instruction, judging the effect of evaluation on the learning process itself, and then assessing the usage of evaluation results. He posits that evaluation is more likely to have effect when (1) important and relatively stable characteristics are measured, (2) the data are used for important decisions at major disjunctions, (3) the results become part of the public record, and (4) they are used to judge the effectiveness of teachers, schools, or school systems.

Stake and Denny (1969) make the cogent observation that evaluation is not "a search for cause and effect, an inventory of present status, or a prediction of future success (p. 370)." It is something of all three but only if it contributes to understanding substance, function, and worth. Evaluation may have many purposes but should always attempt "to describe something and to indicate its perceived merits and shortcomings". They also point out that the successful application of evaluation methodology is dependent upon people and that one of the most urgent needs is for more effective recruitment and training of evaluators. This need is not supplied by graduate programs in

tests-and-measurements or in research-design-and-analysis. To realize the potential of evaluation, different concepts and techniques are needed. Other needs are better means of representing educational goals and priorities, reliable techniques for assessing instruction and learning materials, and simple instruments for gathering opinions and value judgments. Needed most is the development of formal procedures for processing evaluative information and drawing inferences that warrant the confidence that is placed in them.

A SHARPENING FOCUS

The American Educational Research Association's recurrent interest in educational evaluation is indicated by the April 1970 issue of Review of Educational Research and the monograph series it has launched on the subject of curriculum evaluation. The growing concern for evaluation theory, research, and practice is reflected quite strongly in the issue of RER and several of the conceptual shifts in role of evaluation are pinpointed quite well. Cohen (1970) stressed the radical change in the purpose of evaluation as the result of the national thrust against poverty and discrimination. The large-scale programs of the federal government meant that educators were now involved in social action programs and not merely education as they were accustomed to it. This meant further that the whole business of evaluation research had been thrusted in the politics. This turn of events was a source of confusion to many educators who were not prepared for the shift in national goals and expectations.

The AERA Committee on Curriculum Evaluation was established in 1965 and commissioned to develop guidelines for the development and revision of educational curricula. The Committee has proceeded on the belief that contemporary testing and measurement procedures are inadequate and that special

observation, data-gathering, and decision-making techniques are required. To encourage the development of a more adequate theory and rationale for the use of such techniques, the Committee has issued several monographs in the series that open up discussion in a commendable fashion.

Scriven's (1967) distinction between formative and summative evaluation has been widely adopted. He describes the former as the efforts of educators to improve curricula during their development and the latter as the final or conclusive evaluation that would be made after termination of a course, program, or project. Formative evaluation is, of course, intermediate in the sense that it permits adjustments, modifications, and improvements while the program or project is still in a state of development. The focus is on deficiencies and weaknesses that should be eliminated or on strengths that should be emphasized. In summative evaluation the focus is on judgments or over-all conclusions that would assess the terminal worth or value of the program. The finality or conclusiveness of the evaluation is, needless to say, still a matter of relative degree but the implication is that summative evaluation borders on the final word that may be had on the subject under investigation.

Scriven's notion of formative and summative evaluation is comparable in many ways to the distinctions that Cronbach and Suppes (1969) made elsewhere between decision-oriented and conclusion-oriented research. Traditional methods of science are said by the latter to be directed to the derivation of general conclusions and not to the solution of immediate or practical problems. Conclusion-oriented research is directed by the researcher's own commitments and intellectual interests. He raises questions as the research proceeds and is expected to follow leads that may prove more promising than his initial or primary hypothesis. It is expected that his final results

will contribute to the large body of scientific knowledge in the particular area in which he is working.

Decision-oriented research is conducted for the specific purpose of facilitating decisions that must be made in the immediate future. The purpose of the research is not determined so much by the researcher's interests and commitments as it is by the needs and dictates of persons administratively responsible for the decision. Time constraints and other restrictions are placed on the researcher's efforts and he is not free to follow up side-leads or by-paths that might prove more interesting than his original charge. When the research is completed, the researcher may or may not be free to publish his findings in a scientific or professional journal. His audience will more likely be the administrative, legislative, or public authority for which he did the research and his responsibility will not be limited to publication of his research findings. To the contrary, he may spend months explaining his findings to groups of critics, adversaries, and others who view his research findings with suspicion, if not outright hostility.

EVALUATION AND DECISION MAKING

The most ambitious approach to program evaluation has been made by the Phi Delta Kappa Study Committee on Evaluation (1971). This Committee makes a direct attack on the notion that evaluation is "a healthy science, operating from a base of well-established theory and methodology, and with obvious benefit for all (p. 314)." Evaluation is defined by the Committee as "the process of delineating, obtaining, and providing useful information for judging decision alternatives (p. 40)."

The central concept in evaluation, as presented by the Committee, is decision. The decision process is said to consist of the phases of awareness,

design, choice, and action. Decisions are said to be made in a setting that is either metamorphic, homeostatic, incremental, or neomobilistic. Decision models that are applicable for the decision settings are the synoptic, disjointed incremental, and planned change models. The types of decisions called for may be either planning, structuring, implementing, or recycling decisions.

To mesh the concept of decision with educational evaluation, four basic kinds of evaluation are identified as context, input, process, and product. A flow chart, depicting the basic elements of decision settings, models, and types, are believed by the committee to present a total evaluation model for use by those who would approach evaluation systematically. The extent to which they actually synthesize a new definition and methodology of evaluation is not immediately obvious. Yet, given the criteria set by the Committee, the problems involved in developing an adequate theory of evaluation, and the unresolved issues that remain, the PDK Committee has written a book that anyone taking evaluation seriously should not ignore. They have definitely designed wheels that need not be re-invented. Yet, the complexity of the evaluation model and the ambitiousness of their efforts to treat all forms of evaluation may have resulted in a more detailed model than most persons concerned with evaluation will desire. There is a criterion of fruitfulness that remains to be fully tested.

EVALUATION RESEARCH

The emergence of program evaluation as a research specialty is depicted more fully by the work of Edward Suchman (1967), Rossi and Williams (1972), and Carol Weiss (1972). Others could be added to the list, but the three will suffice.

Weiss' text is based on the theme that evaluation uses the tools and methods of social research in a context of social action where traditional research methods often get bogged down. As social research, evaluation demands an imposing array of talent. Not only are research skills required but also political, social, and relational skills that enable the evaluation researcher to utilize his skills for the purpose of improving the rationality of social policy. The purpose of evaluation research, as stated by Weiss, is to "measure the effects of a program against the goals it sets out to accomplish as a means of contributing to subsequent decision making about the program and improving future programming (p. 4)."

Rossi and Williams believe evaluation research is an input to policy analysis. Its purpose is to provide evidence of the effectiveness of existing programs and proposed alternatives as a basic guide to decision making. As such, it involves both the evaluation of program outcomes and field experiments that would test the merit of new ideas with programmatic implications. Both forms of research should be distinguished from onsite monitoring and should help improve the social policy process. Rossi and Williams leave no doubts that policy-oriented research is indeed political. Not only are the problems methodological but they are organizational and bureaucratic. They state explicitly that the role of the policy analyst and the evaluation researcher is basically forensic. It is their task to gather evidence that will persuade and influence those who must make decisions concerning the program under study. The continuance, modification, or cancellation of the program may depend directly on their skill in doing so.

As Harold Orlans (1973) has pointed out, the term program evaluation occurred only rarely in research literature until the domestic programs of the federal government proliferated in the sixties. Now there seems to be a

consensus among social scientists familiar with governmental affairs that a much needed form of research is one that would evaluate the effectiveness of governmental programs.

Yet, Orlans is dubious that social scientists are either willing or able to provide the kind of evaluations that are so obviously needed. There are basic incompatibilities between research as it has been taught in most graduate schools and evaluation research as it is needed in the arena of public policy. Orlans writes that, "evaluation also faces a set of interlinked moral and practical problems that impair its objectivity, insight, and candor (p. 123)." The successful evaluation researcher will be hard put to supply the adaptability, ingenuity, and resourcefulness that can produce an informed, careful, and relatively impartial research report. The motives to criticize whatever findings or conclusions he reaches will be numerous and uninhibited.

A CONCEPTUAL FRAMEWORK FOR PROGRAM EVALUATION

An emerging research specialty may or may not be reflected in the conceptual shifts seen in educational evaluation by the various committees of NSSE, AERA, and PDK. The evidence does suggest, however, that a more sophisticated form of program evaluation is necessary and that it may not be readily provided by persons trained in conventional research methodology. The difficulties stem from differences in approach, procedure, and expectations; they are both conceptual and methodological in nature.

The conceptual shifts that are involved in the changing demands placed upon educational evaluation may be summarized as: (1) an increasing emphasis on the use of tests and measurements for feedback and direct improvement as opposed to selection and placement of students, (2) an increasing advocacy

of developmental, applied, practical, problem-solving research as opposed to basic, long-term, theory-based research, (3) a lessening of interest in conventional experimental research as opposed to research that is administratively or action oriented, and (4) a growing interest in decision-oriented research as opposed to conclusion-oriented research. For the evaluation of academic programs, these shifts would suggest that a heavier emphasis should be placed on formative, as opposed to summative, evaluation and that the results of evaluation studies will be used increasingly for policy decisions that will shape or determine program substance, form, and content. The extent to which evaluation research will be cast in the mold of operations research or institutional research remains to be seen. Also in question is the extent to which evaluation researchers can maintain scientific objectivity and detachment as these values are traditionally conceived. The degree of active involvement that is implied in some evaluation efforts require a significant alteration of the researcher's professional role as it has previously been viewed.

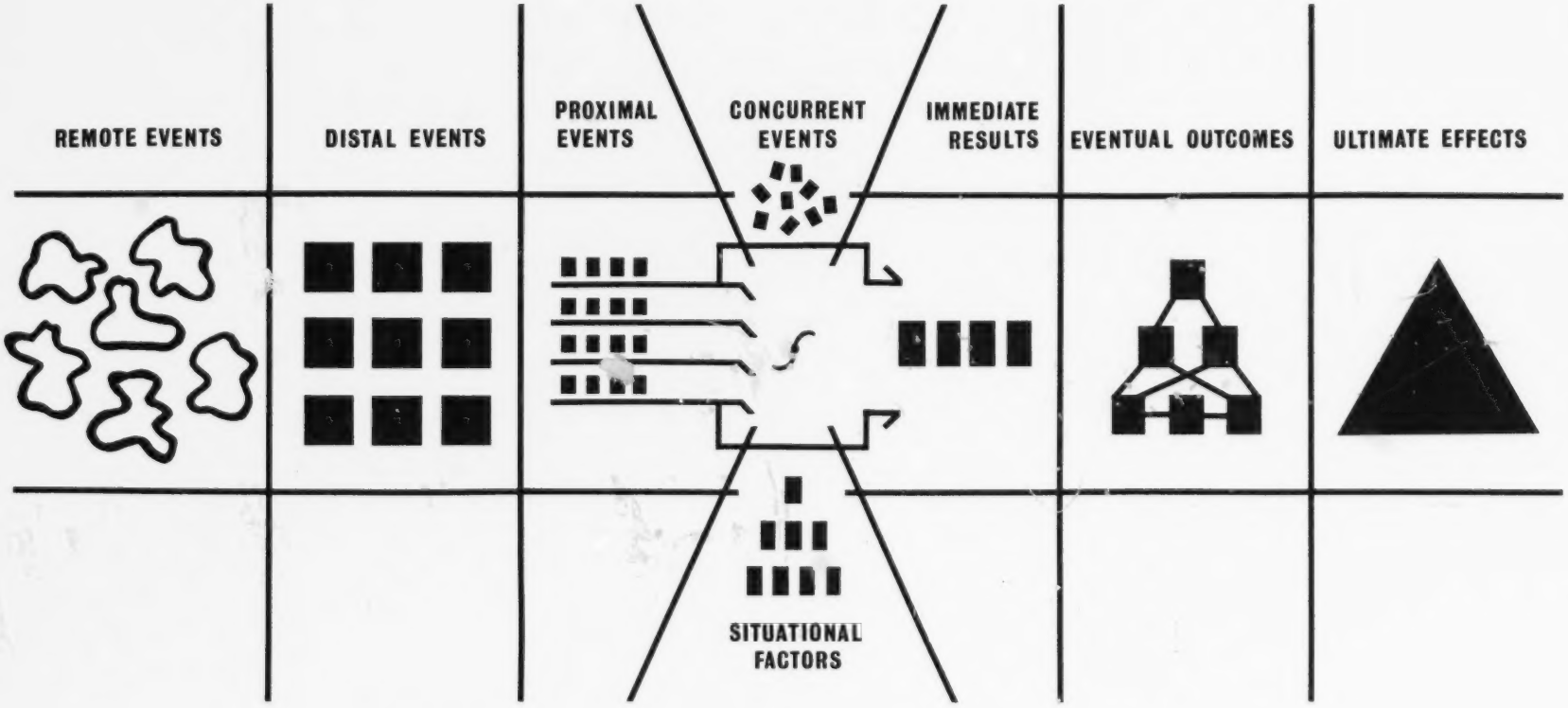
An adequate conceptual framework in which to view program evaluation would seem most desirable. The model developed by the PDK Committee is highly suggestive and should help conceptualize the problems and issues of evaluation in a more meaningful fashion. There is a liability of intricacy, however, that may prevent the model from having the influence it deserves. Persons asked to conduct evaluation studies will continue to be drawn from traditional fields and may find the model to be needlessly complex for their immediate purposes. As a result, many beneficial effects will not be gained.

To provide something of a conceptual framework for program evaluation in higher education, a provisional effort has been made in Figure 1. No effort is made to treat the framework as an operational model or to imply

Figure 1.

A CONCEPTUAL FRAMEWORK FOR PROGRAM EVALUATION

(adapted from E. Brunswik, "The conceptual focus of some psychological systems," Journal of Unified Science, 1939, 8, 36-49)



that it has elaborate structural features. It is strictly a functional representation of the classes of variables that are involved to varying degrees in evaluation efforts. The intent is to provide a perspective from which to view the difficulties, problems, and issues of evaluation. Many efforts at evaluation research would seem to suffer from too narrow a focus on either the process being investigated or on the tangible products or outcomes of that process. The conceptual framework should help demonstrate that there is an optimal distance from which to view most educational programs and projects; there is risk involved in standing either too close or too far back. It is possible both to magnify process variables out of proportion to their importance and to blur their features to such an extent that the process acquires blackbox status.

The conceptual framework or design has been borrowed from Brunswik's writings on probabilistic functionalism and should suggest the advantages of both probabilism and functional analysis. A probabilistic approach should help avoid a narrowly construed determinism that sets the researcher in search of one-to-one relationships between the several classes of variables. The time perspective should help disabuse any researcher of the notion that his efforts should serve to link remote events to ultimate effects. These two classes of variables are provided primarily for boundary purposes and should help discourage the excessive ambitiousness that some evaluation studies have been susceptible to in the past. The functional analysis of distal events and eventual outcomes should give any researcher as much as he could hope to handle.

It should be obvious from the conceptual design that a choice of variables is crucial to the process of evaluation. Proximal events will be identified as those variables that have been traditionally classified as

independent or experimental variables while concurrent events will be quickly recognized as extraneous variables that must be controlled if functional relationships are to be clearly established. Situational factors may be seen as a special class of variables in that they have an anchoring effect on the decisions or conclusions that may be derived from the evaluation. This is to say that virtually all evaluation results will have some degree of specificity that cannot be generalized to other situations and circumstances.

Immediate results will be seen as comparable to the dependent or behavioral variables that are identified in conventional experimental designs. These will be the variables in which the researcher will be most often interested because they are more accessible, more tangible, and more easily communicated to others in decision making positions. If, however, there is a serious concern with eventual outcomes, the evaluation study must be properly executed to that end. The distinction between immediate results and eventual outcomes will depend, of course, on the researcher's own time perspective. The distinctions are broadly implied, however, and if eventual outcomes are to be treated as dependent variables, the rationale and procedures for doing so should be specified. On the other hand, if immediate results are to be treated as proxies for eventual outcomes, the reasons for doing so should be explicit.

Distinctions between types, levels, and forms of evaluation are not inherent in the conceptual design but should be facilitated to some extent by its consideration. An obvious inference is that all classes of variables need not be included in a particular evaluation study. If the conceptual framework provides any help at all, it should convey the foolishness of efforts or notions that we can evaluate academic programs in a total or comprehensive manner. The effectiveness of any evaluation effort will always depend on the

researcher's ability to identify, analyze and interpret those variables that bear most directly on the process, outcomes, or events he has chosen to study. The choice of variables, therefore, should always be explicit and as consistent as possible with the overall objectives of the evaluation study. For example, the choice of variables in organizing and implementing a system of program evaluation should reflect the difference between a monitoring system and a means for individual certification. Just as testing practices have been dominated by reliability theory as a means of providing highly stable measures of individual performance, so can evaluation efforts be derailed by the mistaken notion that the purpose is to certify individual competence. The two approaches are different and require different techniques and procedures.

In closing, it is well to mention that the conceptual framework should incorporate without undue difficulty the major advantages and implications of the schema, models, and approaches discussed previously. The primary insistence of the conceptual design is that systematic efforts at program evaluation should constitute a functional analysis of the variables that are involved and that the nature of the relationships derived or established will be probabilistic rather than deterministic. It is believed that such a conceptual approach can help avoid many of the difficulties that have been previously involved in our efforts to evaluate academic programs. The complexities of most programs in higher education and the subtlety of influential variables would strongly suggest that as urgent as evaluation is, it will not be easy to conduct the kind of evaluation studies that are needed. The methodological difficulties are staggering but they are not insurmountable. There is every reason to believe that our success or failure in program evaluation will depend primarily on our solution of the conceptual problems that are involved.

References

Averch, Harvey A., Stephen J. Carroll, Theodore S. Donaldson, Herbert J. Kiesling, and John Pincus. How Effective Is Schooling? A Critical Review and Synthesis of Research Findings. Santa Monica, Calif.: RAND Corp., 1972.

Bloom, Benjamin S. "Some Theoretical Issues Relating to Educational Evaluation." In Ralph W. Tyler (Ed.), Educational Evaluation: New Roles, New Means. Chicago: University of Chicago Press, 1969, pp. 26-50.

Campbell, Donald T. and Julian C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1963. (Reprinted from Handbook of Research on Teaching).

Carnegie Commission on Higher Education. Reform on Campus: Changing Students, Changing Academic Programs. New York: McGraw-Hill, 1972.

Cohen, David K. "Politics and Research: Evaluation of Social Action Programs in Education." Review of Educational Research, 1970, 40, 213-238.

Coleman, James S., et al. Equality of Educational Opportunity. Washington, D. C.: U. S. Government Printing Office, 1966.

Commission on Nontraditional Study. Design for Diversity. San Francisco: Jossey-Bass, 1973.

Cronbach, Lee J. "Course Improvement Through Evaluation." Teacher's College Record, 1963, 64, 672-682.

Cronbach, Lee J. and Patrick Suppes (Eds.) Research for Tomorrow's Schools: Disciplined Inquiry for Education. New York: National Academy of Education, 1969.

Denny, Terry (Issue Editor). "Educational Evaluation." Review of Educational Research, 1970, 40(2).

Jencks, Christopher, et al. Inequality: A Reassessment of the Effect of Family and Schooling in America. New York: Basic Books, 1972.

Merwin, Jack C. "Historical Review of Changing Concepts of Evaluation." In Ralph W. Tyler (Ed.), Educational Evaluation: New Roles, New Means. Chicago: University of Chicago Press, 1969, pp. 6-25.

Orlans, Harold. Contracting for Knowledge. San Francisco: Jossey-Bass, 1973.

Provus, Malcolm. "Evaluation of Ongoing Programs in the Public School System." In Ralph W. Tyler (Ed.), Educational Evaluation: New Roles, New Means. Chicago: University of Chicago Press, 1969, pp. 242-283.

Rossi, Peter H. and Walter Williams. Evaluating Social Programs: Theory, Practice, and Politics. New York: Seminar Press, 1972.

Scriven, Michael. "The Methodology of Evaluation." Perspectives of Curriculum Evaluation. Chicago: Rand McNally, 1967, pp. 39-83.

Stake, Robert E. and Terry Denny. "Needed Concepts and Techniques for Utilizing More Fully the Potential of Evaluation." In Ralph W. Tyler (Ed.), Educational Evaluation: New Roles, New Means. Chicago: University of Chicago Press, 1969, pp. 370-390.

Stufflebeam, Daniel L., Walter J. Foley, William J. Gephart, Egon G. Guba, Robert L. Hammond, Howard O. Merriman, and Malcolm M. Provus (PDK National Study Committee on Evaluation). Educational Evaluation and Decision Making. Itasca, Ill.: F. E. Peacock, 1971.

Suchman, Edward A. Evaluative Research: Principles and Practice In Public Service and Social Action Programs. New York: Russell Sage Foundation, 1967.

Tyler, Ralph W., Robert M. Gagné, and Michael Scriven. Perspectives of Curriculum Evaluation. Chicago: Rand McNally, 1967.

Tyler, Ralph W. (Ed.) Educational Evaluation: New Roles, New Means. The Sixty-eighth Yearbook of the National Society for the Study of Education, Part II, Chicago: University of Chicago Press, 1969.

Weiss, Carol H. Evaluation Research: Methods of Assessing Program Effectiveness. Englewood Cliffs: Prentice-Hall, 1972.